Homework 3 - ECON 21130

T. Lamadon

Part 1 (25 points)

We consider a measurement system for a random variable of interest X which is unobserved to the econometrician. Instead we observe two variables Y_1, Y_2 related to X in the following way:

$$Y_{i1} = X_i + \epsilon_{i1}$$
$$Y_{i2} = X_i + \epsilon_{i2}$$

and we assume that $\mathbb{E}(X) = \mu$, $\text{Var}(X) = \sigma_x^2$, $X \perp \epsilon_1 \perp \epsilon_2$, $\mathbb{E}(\epsilon_1) = \mathbb{E}(\epsilon_2) = 0$ and $\text{Var}(\epsilon_1) = \sigma_1^2, \text{Var}(\epsilon_2) = \sigma_2^2$. Given a set of independent draws $(Y_{i1}, Y_{i2})_{i=1..n}$ from our model, we consider the following estimator for μ:

$$\mu_n = \alpha \frac{1}{n} \sum_{i=1}^{n} Y_{1i} + (1 - \alpha) \frac{1}{n} \sum_{i=1}^{n} Y_{2i}$$

1. (5 points) Show that μ_n is an unbiased estimator of μ.

2. (5 points) Show that μ_n is a consistent estimator of μ.

3. (5 points) Compute the variance of μ_n as a function of α and the other parameters.

4. (6 points) Find α as a function of σ_1, σ_2 that minimizes the variance. Interpret the results.

5. (4 points) We assumed through this question that $\epsilon_1, \epsilon_2, X$ were independent. Could we find a weaker assumption and keep the unbiasedness?

Part 2 (25 points)

We keep the same model as in Part 1: $Y_{ik} = X + \epsilon_{ik}$ with $\epsilon_1 \perp \epsilon_2$, $\mathbb{E}(\epsilon_k) = 0$ and $\text{Var}(\epsilon_k) = \sigma_k^2$. In this part 2, **we further assume** that the distributions of ϵ_k are normal $\epsilon_k \sim \mathcal{N}(0, \sigma_k^2)$, and that $\text{Var}(X) = 0$ such that $X = \mu$ is a constant. Under these assumptions Y_1, Y_2 are normally distributed as well.

1. (5 points) Show that under the current assumptions ($X = \mu$, $\sigma_x = 0$) we have that $Y_1 \perp Y_2$. If, as in Part 1, we had that $\sigma_x > 0$ what would $\text{Cov}(Y_1, Y_2)$ be?

2. (8 points) Given that Y_1, Y_2 are independent here and the underlying distributional assumption, compute the likelihood and the log-likelihood of observing $(Y_{i1}, Y_{i2})_{i=1}^n$ given parameters μ, σ_1, σ_2.

3. (7 points) Finally, solve for the maximum likelihood estimator for μ (consider here that the variances σ_1, σ_2 are known). Explain your steps.

4. (5 points) Compare and contrast this estimator to the optimal weighting found in question (3), part 1.

Part 3 - True, False or Uncertain (25 points)

For each of the following questions, please answer whether the statement is True, False or Uncertain, and then justify your answer with a short paragraph, a proof, or a counterexample as appropriate. No credit will be given for answers without any justification.

1. (5 points) Let U be a uniform random variable on $[0, 1]$, then $\mathbb{E}[U^\gamma] = \frac{1}{\gamma}$.

2. (5 points) Consider the ML estimator θ_n for a parameter θ. This estimator is always unbiased for θ.

3. (5 points) All econometric estimators achieve $\sqrt{n}$ convergence to a normal distribution, i.e.

$$\sqrt{n}\left(\hat{\theta}_n - \theta\right) \xrightarrow{\text{d}} \mathcal{N}(0, \sigma^2)$$

4. (5 points) Consider the data generating process to be $D_i = \mathbf{1}\left[X_i'\beta + \epsilon_i > 0\right]$ with ϵ_i normally distributed and independent across individuals, then it is the case that

$$Pr\{D_i = 1 \mid X_i\} = \frac{1}{1 + \exp(-X_i'\beta)}$$

5. (5 points) Let $X_1, \ldots, X_n$ denote the number of Bernoulli trials with p until the first success happens, which can be modeled as a geometric distribution parameterized by p. Therefore, if k is the number of trials, then

$$\Pr\{X_i = k\} = (1-p)^{k-1}p$$

This implies that $\mathbb{E}[X_i] = \frac{1}{p}$.

Part 4 - A DIY ML Estimator (25 points)

You are asked to construct the ML algorithm from scratch for a simple model. Follow the sequence of steps provided to get full credit for this question. DO NOT use any pre-packaged, high-level econometric functions to work through this question, as your grade will reflect your understanding and execution of the ML algorithm without the aid of Stata-like commands. This will prove useful later in your career when you develop and estimate structural models!

1. Simulate 1000 independent draws (indexed by i) of an unobserved variable ϵ_i drawn from $\mathcal{N}(0,4)$. Similarly, generate data for the observed variable X_i from $\mathcal{N}(0,1)$. Finally, construct

$$Y_i = X_i\beta + \epsilon_i$$

   Take $\beta = 1.5$. You can delete ϵ_i from your data now, to replicate the conditions under which you would normally perform such regressions.

Now, we treat β and σ_ϵ^2 as the parameters of interest, which can be recovered using MLE from the data with observations for Y_i, X_i.

2. Construct a function that computes the log-likelihood associated with the sample given inputs for β and σ_ϵ^2.

3. Using the function above, attempt to maximize the log-likelihood (or minimize the negative log-likelihood) to recover the ML estimates for β and σ_ϵ^2. You may use an optimizer function or package.

4. (Bonus) Compute the standard errors for your estimators.